An Application of Computational Algebraic Geometry in Signal Processing

John B. Little

Department of Mathematics and Computer Science

College of the Holy Cross

Worcester, MA 01610

email: little@mathcs.holycross.edu

Ruth Michler Memorial Conference

U.S. Naval Academy

October 28, 2001

Outline of Talk

1. "Review" of digital filters

2. A filter design problem

3. Examples

4. A solution method, based on algebraic geometry

Digital Signals

A *digital* signal is a quantized function of a *discrete* variable.

- only have signal values at $n\Delta t$, n an integer, Δt the *sampling interval*

- signal values are elements of a discrete set.

For simplicity, we ignore quantization effects and study *discrete-time, complex-valued* signals, $x[n]$. Real-world signals are "time-limited": $x[n] = 0$ for all $|n| \geq N$, some N. Could also consider $x[n] \in \ell^2$ (finite energy condition).

Filters

Signal processing operations can be described mathematically by means of functions Γ:

$$\Gamma : \{\text{signals}\} \to \{\text{signals}\}.$$

We mainly consider operations with the following properties

1. (Linearity) If $z[n] = cx[n] + y[n]$, and $w = \Gamma(z), u = \Gamma(x), v = \Gamma(y)$, then
$$w[n] = cu[n] + v[n].$$

2. (Shift-invariance) If $y[n] = x[n + k]$ for all n, k fixed, and the outputs are $w = \Gamma(x)$, $z = \Gamma(y)$, then
$$z[n] = \Gamma(y)[n] = \Gamma(x)[n + k] = w[n + k]$$

A Γ satisfying both of these properties is said to be a (linear, shift-invariant) *filter*.

A "Moving Average" Filter

For any input signal $x[n]$, define

$$\begin{aligned} y[n] &= \Gamma(x)[n] \\ &= \left(\sum_{k=0}^{4} x[n-k] \right) /5 \end{aligned}$$

If $x[n] = \cos(n\pi/2)$,

$$\cdots, 1, 0, -1, 0, 1, 0, -1, 0, 1, \cdots$$

for instance, then $\Gamma(x)[n] =$

$$\cdots, 0, 1/5, 0, -1/5, 0, 1/5, 0, -1/5, 0, \cdots$$

Linearity and shift-invariance properties hold, so this defines a filter.

Impulse Response of a Filter

A nice result about filters: Let δ be

$$\cdots, 0, 0, 1, 0, 0, \cdots$$

(1 at time $n = 0$) – called a *unit impulse* input. $\Gamma(\delta)$ is called the *impulse response* of the filter.

If we know $h = \Gamma(\delta)$, then we also know the output of the filter on any other signal x. Reason: We can write

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k]$$

Hence by linearity and shift-invariance if $y = \Gamma(x)$, then

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k].$$

The Transfer Function

Package $x[n], y[n], h[n]$ by their "z-transforms" X, Y, Z, e.g.

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n}$$

Coefficient of z^{-n} in $H(z)X(z)$ is the discrete convolution:

$$\sum_{k=-\infty}^{\infty} h[k]x[n-k],$$

same as $y[n]$ (!)

$\Rightarrow Y(z) = H(z)X(z)$. $H(z)$ is called the *transfer function* of the filter.

Note:

$$H(e^{i\omega}) = \sum_{n=-\infty}^{\infty} h[n]e^{-in\omega}$$

is the (discrete-time) Fourier transform of h, so $H(z)$ also determines the *frequency response* characteristics of the filter on input signals.

Filter Design

Goal is to design filters to perform various operations on signals, by designing $H(z)$.

- Low-pass filters: remove high-frequency components of signals – typically *smooth out* or *blur* signals (can be used to remove high-frequency "noise")

- High-pass filters: remove low-frequency components – typically focus on fine details, rapid changes (can be used to detect features)

Typical problems: design a *low-pass* or *high pass* filter to meet certain specifications.

Selesnick-Burrus Design Paradigm

reference: Selesnick and Burrus, *IEEE Trans. Circuits and Systems II*, v. 45(1), January 1998.

Given an FIR low-pass filter

$$H(z) = \sum_{n=0}^{N} h[n]z^{-n},$$

let $F(\omega) = |H(e^{i\omega})|^2$, $G(\omega)=$ "group delay". If $K + L + M + 1 \leq N$, can determine $h[n]$ to make:

- $F^{(2i)}(0) = 0$, $i = 1, \ldots, M$,
- $G^{(2j)}(0) = 0$, $j = 1, \ldots, L$, and
- $(1 + z)^K | H(z)$

Form of The Equations in $m[i]$

In terms of the filter coefficient *moments*,

$$m_k = \sum_{n=0}^{N} n^k h[n],$$

the corresponding equations are:

- Flatness equations (quadratic in the $m[i]$):

$$\binom{2i}{i} m_i^2 + 2 \sum_{\ell=0}^{i-1} \binom{2i}{\ell} (-1)^{i+\ell} m_\ell m_{2i-\ell}$$

$= 0$, $i = 1, \ldots, M$.

$$\sum_{\ell=0}^{j} (1 - \frac{2\ell}{2j+1}) \binom{2j+1}{\ell} (-1)^\ell m_\ell m_{2j+1-\ell}$$

$= 0$, $j = 1, \ldots, L$.

- Zeros at $z = -1$: remainder of $H(z)$ on division by $(1+z)^K$ is zero – K linear equations on m_i

First Simplification

This looks like an underdetermined system.

But can express m_k for $k \geq N$ in terms of $m_0, \ldots, m_{N-1}$ by solving linear equations, and substitute.

Normalization: $m_0 = 1$

Get $L + M + K = N - 1$ equations in $N - 1$ variables $m_1, \ldots, m_{N-1}$.

Expect only finitely many solutions. (Mainly interested in real ones.)

Here is an example.

The Case $L = 1, M = 5, K = 1$

The quadratic equations:

$$
\begin{aligned}
&2m_1^2 - 2m_2,\\
&6m_2^2 + 2m_4 - 8m_1m_3,\\
&20m_3^2 - 2m_6 + 12m_1m_5 - 30m_2m_4,\\
&70m_4^2 + 2m_8 - 16m_1m_7 + 56m_2m_6\\
&\quad -112m_3m_5,\\
&252m_5^2 - 2m_{10} + 20m_1m_9 - 90m_2m_8\\
&\quad +240m_3m_7 - 420m_4m_6,\\
&-m_3 + m_1m_2
\end{aligned}
$$

Example, cont.

The linear equations:

$$
\begin{aligned}
&-315 + 14496m_1 + 23912m_3 - 9310m_4 \\
&\quad +8m_7 - 196m_6 + 1904m_5 - 30184m_2, \\
&2m_8 - 728m_6 + 9408m_5 - 51632m_4 + \\
&\quad 141120m_3 - 185152m_2 + 91392m_1 - 2205, \\
&4m_9 - 17052m_6 + 247380m_5 - 1445010m_4 \\
&\quad +4105160m_3 - 5529048m_2 + 2784096m_1 \\
&\quad -72765, \\
&m_{10} - 43407m_6 + 670320m_5 - 4070200m_4 \\
&\quad +11869200m_3 - 16288944m_2 + 8326080m_1 \\
&\quad -231525
\end{aligned}
$$

Example, cont.

- The *lex* Gröbner basis with $m_{10} > m_9 > \cdots > m_1$ is in generic "Shape Lemma" form, with univariate polynomial of degree 16 in m_1.

- Using numerical root-finding, we find 6 approximate real roots: $m_1 \doteq .04470426799$, 1.233505559, 2.558981682, 4.441018318, 5.766494441, 6.955295732.

- The 6 real roots occur in pairs: $r, 7 - r$

- Up to time reversal, this gives just 3 distinct filters satisfying the Selesnick-Burrus conditions in this case.

Some Notation

We will use the following notation for the ideals generated by the polynomials from the Selesnick-Burrus equations:

- $J_{L,M}$ is the ideal generated by the $L+M$ quadratic flatness equations (doesn't depend on K)

- $I_{K,L,M}$ is the ideal generated by these, plus the K linear equations from the zeroes at $z=-1$, plus the equations expressing the $m_N, \ldots, m_{2M}$ in terms of $m_0, \ldots, m_{N-1}$.

“Easy” and “Hard” Cases

- The *easier* cases are
$$\left\lfloor \frac{M-1}{2} \right\rfloor \le L \le M$$
All variables except m_1 can be eliminated using linear algebra techniques.

- ("Region II") The *harder* cases are
$$0 \le L \le \left\lfloor \frac{M-1}{2} \right\rfloor - 1$$

Example: if $L = 0$, there are 2^M solutions of the design equations in these cases(!) (One way to see this: the generators for $I_{K,0,M}$ are already “almost” a Gröbner basis for $I_{K,0,M}$ with respect to a *grevlex* order with $m_{N-1} > m_{N-2} > \cdots > m_1$.)

Problems

- Understand the Region II cases:

  1. for given K, L, M, how many solutions?

  2. how many are real?

  3. how many yield monotone decreasing square magnitude response $|H(e^{i\omega})|^2$?

  4. etc.

- Develop an efficient method to solve the S-B equations in this case

This is where some insight from algebraic geometry is useful!

The Affine Variety $\mathbf{V}(J_{L,M})$

Of course, nothing special about varieties defined by quadrics, but the S-B equations have a *very particular* form:

- $\mathbf{V}(J_{L,M})$ is smooth of dimension $M - L$

- The S-B quadrics are homogeneous if

  $$\text{weight}(m_i) = i,$$

  hence $\mathbf{V}(J_{L,M})$ contains a rational normal curve passing through each point

- In fact, $\mathbf{V}(J_{L,M})$ is the image of an $(M-L)$-fold *rational normal scroll* under a certain "upper-triangular" automorphism of $\mathbf{A}^{2M}$.

- The SB equations define a linear section of this "warped scroll".

The Scroll: $K = 1, L = 1, M = 5$

Here is the *lex* Gröbner basis for $J_{1,5}$ (with $m_{10} > m_9 > \cdots > m_1$):

$$
\begin{aligned}
\{ & m_2 - m_1^2, \quad m_3 - m_1^3, \quad m_4 - m_1^4, \\
& m_6 - 6m_1m_5 + 5m_1^6, \\
& m_8 - 8m_1m_7 + 112m_1^3m_5 - 105m_1^8, \\
& m_{10} - 10m_1m_9 + 240m_1^3m_7 - 126m_5^2 \\
& \quad -3780m_1^5m_5 + 3675m_1^{10}\}
\end{aligned}
$$

The related ideal

$$
\begin{aligned}
J' \; = \; & \langle m_2 - m_1^2, m_3 - m_1^3, m_4 - m_1^4, m_6 - 6m_1m_5, \\
& m_8 - 8m_1m_7, m_{10} - 10m_1m_9 \rangle
\end{aligned}
$$

is equal to the ideal generated by the 2×2 minors of

$$
\begin{pmatrix}
m_1 & m_2 & m_3 & m_4 & m_6 & m_8 & m_{10} \\
1 & m_1 & m_2 & m_3 & 6m_5 & 8m_7 & 10m_9
\end{pmatrix}
$$

The Scroll, cont.

Hence, $S = \mathbf{V}(J')$ is an affine *4-fold rational normal scroll* = union of $\mathbf{A}^3$'s spanned by related points on a rational normal curve of degree 4 and 3 lines.

$\mathbf{V}(J_{1,5})$ is $\alpha(S)$ for the automorphism of $\mathbf{A}^{10}$ defined by $\alpha(m_1, \ldots, m_{10}) =$

$$\begin{aligned}(m_1, \ldots, m_5, m_6 + 5m_1^6, m_7,\\ m_8 + 112m_5m_1^3 - 105m_1^8, m_9,\\ m_{10} + 240m_1^3m_7 - 126m_5^2\\ -3780m_1^5m_5 + 3675m_1^{10})\end{aligned}$$

$\Rightarrow$ $\mathbf{V}(J_{1,5})$ has a very nice *parametrization* – can be seen from the *lex* Gröbner basis.

The parametrization, L = K = 1, M = 5

$$
\begin{aligned}
m_1 &= t \\
m_2 &= t^2 \\
m_3 &= t^3 \\
m_4 &= t^4 \\
m_5 &= a \\
m_6 &= 6at - 5t^6 \\
m_7 &= b \\
m_8 &= 8bt - 112at^3 + 105t^8 \\
m_9 &= c \\
m_{10} &= 10ct - 240bt^3 + 126a^2 - 3780at^5 \\
&\quad -3675t^{10}\}
\end{aligned}
$$

Similar results hold for all $\mathbf{V}(J_{L,M})$.

The "Best" Solution Method

To solve the SB equations,

- The *lex* GB of $J_{L,M}$ with $m_{N-1} > m_{N-2} > \cdots > m_1$ gives a similar parametrization for $V(J_{L,M})$ for all L, M.

- Substituting from this parametrization into the linear equations gives a new system with quite low degrees in all variables except m_1

- Elimination to a univariate polynomial in m_1 can be done by either

  – sparse resultants,

  – or pairwise ordinary (Sylvester) resultants

  The actual computation is fast (but there are always extraneous factors in the result).

Some Results in Region II

$$\#\text{sols} = \dim \mathbf{C}[m_1, \ldots, m_{2M}]/I_{1,L,M}$$

in Region II ($K = 1$):

M/L	0	1	2	3	4	5
5	32	16				
6	64	26				
7	128	48	24			
8	256	78	38			
9	512	152	66	32		
10	1024		112	50		
11	2048			86	40	
12	4096			142	62	
13	8192				106	48

(Cases $M \leq 7$ check with Selesnick-Burrus. $M \geq 8$ new; blank space below Region II boundary means (?))

A Theoretical Result

Also, using this approach, we can prove:

Theorem 1 *In the cases* $M = 2L + 3$, $L \geq 0$, $K = 1$ *("corners" in Region* II *boundary), the S-B equations always have* $8(L + 1)$ *distinct solutions, all of multiplicity 1. Of these* $2(L + 2)$ *are real* ($L + 2$ *different filters, up to time reversal). Two of these have monotonic square magnitude response.*

Solutions yield square magnitude responses with properties reminiscent of Sturm-Liouville eigenfunctions(?)

Similar techniques should work on other "diagonals" $M = 2L + \ell$ with $\ell \geq 4$ also.

Things That Make These Systems Hard

- High degrees of univariate polynomials

- For large M, L, the univariate polynomial has *gigantic* coefficients.

- A *BIG* numerical analysis problem. Need to use very many decimal digits in floating-point computation to get accurate roots.

- In fact this is the current "bottleneck" in the cases I have worked out.